Differential Object Marking in Structurally Complex Contexts in Spanish: Evidence from Bilingual and Monolingual Processing

Aurora Bel * and Rut Benito

Abstract: This study examines whether Differential Object Marking (DOM) realization and word order in relative clauses (RCs) in Spanish affect processing and interpretation among monolinguals and highly proficient Catalan–Spanish bilinguals. RCs are parallel in Catalan and Spanish, but DOM is much more restricted in Catalan than in Spanish, and, interestingly, the distinction between subject and object RCs relies mainly on the presence/absence of DOM. To examine DOM optionality, we concentrate on the top portion of the animacy scale and test the human/non-human contrast. Exploring these two populations allows us to test whether they resort to different strategies for the following three reasons: (1) bilingualism places an increased burden on memory processes; (2) the partial overlap between both DOM systems might lead to the influence from Catalan into Spanish; and (3) optionality has been proposed to characterize bilingual grammars. Findings from a word-by-word non-cumulative self-paced reading task showed that DOM modulates RC processing. With [+human] obligatorily marked objects, both monolinguals and bilinguals read subject RCs faster than object RCs, suggesting a strategy favoring subject RCs. However, monolinguals solved the interpretation early while processing but bilinguals, despite the more restricted DOM character of Catalan, are sensitive to DOM albeit displaying delayed spill-over effects. With [−human] optionally marked objects, bilinguals performed faster than monolinguals. We suggest that the uneven experience with DOM in Catalan, particularly with the non-standard variety that frequently displays DOM and that our bilinguals also speak in everyday conversations, facilitates bilinguals’ adaptation to the optional marking of non-human objects in Spanish, much in the same manner that they accommodate the presence or absence of DOM with both human and non-human objects in other native language.

Keywords: differential object marking (DOM); relative clauses; word order; Catalan; Spanish bilingualism

1. Introduction

This study explores the processing of Spanish Differential Object Marking (DOM) in interaction with relative word order in relative clauses, investigating Spanish monolingual speakers and Catalan–Spanish speakers who are bilinguals in a bilingual society. DOM refers to the phenomenon where directs objects in a sentence receive different case-markings depending on their inherent properties. Spanish and Catalan are particularly intriguing for this study due to their overlapping options in DOM realization. While Spanish is considered one of the languages that exhibit a broader implementation of DOM, Catalan shows DOM in a reduced subset of objects compared to Spanish. These linguistic peculiarities, coupled with prolonged contact in a bilingual society where the two languages coexist and are spoken, create an ideal environment for crosslinguistic influence (CLI) between the two languages, which can be particularly manifested in the DOM system.

Sustained language contact between languages can lead to crosslinguistic influence (CLI). CLI (Silva-Corvalán 2014) is an effect of a grammatical feature of one language on the other (Jarvis and Pavlenko 2008). In our scenario, the overlap of Catalan and Spanish DOM
systems, as proposed in an expanded interpretation of Hulk and Müller’s (2000) hypothesis, could potentially result in the influence from Catalan into Spanish. Spanish and Catalan are two languages that have been in contact for decades and currently coexist in a bilingual context. These two languages share many phonological, syntactic, and semantic features, as well as numerous lexical items. Linguistic similarity serves as a catalyst for crosslinguistic influence (CLI), prompting researchers to investigate CLI between Catalan and Spanish (Bel and García-Alcaraz 2018; Davidson 2012; Paolieri et al. 2020; Perpiñán 2018; Perpiñán and Soto-Corominas 2021, among others). This phenomenon has led some scholars to propose the existence of a Catalan bilingual variety (Boix and Sanz 2008), particularly in Catalonia, a historically bilingual region where both languages are co-official and commonly used in everyday communication.

This paper focuses on the CLI effects at the intersection of case-marking (specifically, DOM) and the word order involved in subject and object relative clauses (RC) in Spanish. Catalan and Spanish show different extensions of DOM. In Spanish, DOM is obligatory with [+human] objects and optional with [−human] ones (Real Academia Española 2010, p. 2635); inanimate objects do not bear DOM, but these objects are beyond the scope of the current study. Catalan, in its standard variety, has been argued to be virtually a non-DOM language; according to prescriptive grammar, DOM is limited to personal pronouns acting as objects (along with certain constructions at the syntax–pragmatics interface).

DOM marking holds particular significance in complex sentences such as sentences containing RCs. In Spanish RCs, ambiguity may arise when the subject and the object coincide in number and person features, with (present or absent) DOM serving as the disambiguating factor. This crucial aspect motivates our research, as we aim to specifically examine these ambiguous structures. In Catalan, the absence of generalized DOM contributes to broader ambiguity in argument interpretation within RCs. To elucidate the significance of examining RCs, it is noteworthy that, unlike matrix clauses, DOM plays a pivotal role in RCs. In fact, subject and object RCs are structurally ambiguous until the postverbal constituent is reached. If the postverbal constituent is case-marked, the clause is a subject RC (SVO), with the postverbal constituent identified as an object. If it is not marked, then the postverbal constituent is a postponed subject, and the clause becomes an object RC (OVS). To provide clarity, we present an example here that will be further discussed in the introduction to Section 1.3: In a sentence like “La tenista que persigue…” (“The tennis player that chases…”), the role (subject or object) of the subsequent argument depends on the presence or absence of DOM: a “la periodista” vs. “la periodista” (“(DOM) the journalist”). The ambiguity is solved by DOM.

In addition, another source of ambiguity arises from the optional nature of DOM in Spanish with [+animate, −human] objects. Taken together, these aspects, which lead to ambiguity, along with the stricter usage of DOM in standard Catalan, provide an excellent opportunity to examine the interpretation and processing of potentially ambiguous sentences by early Catalan–Spanish-balanced bilingual individuals.

The comparison of subject and object RCs will allow us to ascertain whether they differ in processing complexity, as found in prior research for different languages using varying methodologies (for Spanish, see Betancort 2006; Betancort et al. 2009; Del Río et al. 2012; Sagarra et al. 2017, among others). As mentioned in the preceding paragraph, the involvement of DOM in distinguishing between subject and object RCs renders the comparison between monolingual and bilingual RC processing particularly intriguing. Catalan exhibits a significantly more limited use of DOM compared to Spanish, as previously noted (the standard variety of Catalan does not mark animate objects). Consequently, Catalan speakers must suppress the influence of non-DOM marking from Catalan when discerning and interpreting subject and object RCs in Spanish. This situation may increase cognitive load due to the ongoing need to resolve potential conflicts between the two language systems (Bialystok et al. 2008; Costa et al. 2006; Sorace 2016, among others), potentially resulting in more demanding processing that could manifest in longer reading times. In this respect, the current research is unique in providing information on how monolinguals
and bilinguals exploit these two cues (DOM and word order) jointly or separately, with one being potentially more prominent or efficient in structural processing.

The interplay between case-marking and word order—the one involved in RCs—is not incidental, since its study can create new insights based on the existing knowledge of morphosyntactic processing by conducting innovative research in the interaction between multiple linguistic cues (e.g., Comeaux and McDonald 2018; Henry et al. 2020). The findings of this piece of research can also inform bilingual processing studies.

1.1. Differential Object Marking and Relative Clauses in Spanish and Catalan

Differential Object Marking, which exists in a large variety of languages (Bossong 1991), depends on different morphosyntactic, semantic, and pragmatic features of the direct object, the verb, and the subject, with their contributions varying cross-linguistically (Dalrymple and Nikolaeva 2011). The combination of these features determines whether the object of a transitive verb is marked with a grammatical element or not. DOM realization in Spanish (a morpheme) is regulated by a combination of the Animacy and Definiteness scales (Aissen 2003), as represented in (1), and marks accusative inherent case. These scales, proposed to explain the crosslinguistic tendencies of DOM systems, predict that higher NPs in these scales have a higher probability of being marked than lower NPs. Although specificity is also a feature involved in DOM systems, our focus in this paper is solely on the definite portion of this scale (as we will discuss, our object stimuli always include a definite article).

(1) a. Animacy scale: human > animate > inanimate;
   b. Definiteness scale: personal pronoun > proper name > definite NP > non-definite, specific NP > non-definite, non-specific NP.

In Peninsular Spanish, DOM is mandatory with [+human] objects (as in (2a)) and optional with [+animate, −human] objects (i.e., animals) (as in (2b)). On the contrary, [−animate] direct objects are not marked (as in (2c)) (Torrego 1998; López 2012).

(2) a. La gata que buscó a la niña entró inesperadamente. [+animate, +human]
   ‘The cat that looked for the girl entered unexpectedly.’
   b. La niña que buscó (a) la gata entró inesperadamente. [+animate, −human]
   ‘The girl that looked for (DOM) the cat entered unexpectedly.’
   c. La gata que buscó a la mesa entró inesperadamente. [−animate]
   ‘The cat that looked for the table entered unexpectedly.’

The system is, in fact, much more complex because not all cases follow these general rules (see Fábregas 2013 for further discussion): studies have demonstrated that features, such as definiteness, specificity, affectedness, or telicity, also play a role in the DOM system (Laca 2006). For the present study, only cases in which animacy is involved will be considered. Spanish, as well as Catalan, also mark dative case with ‘a’.

An interesting context in which the opposition between mandatory and optional DOM, depending on animacy, can be found is in RCs where the subject and object share the same number and person features. In subject RCs (SVO), the direct object is compulsorily marked if it is [+animate, +human], as in (3a), but optionally marked if it is [+animate, −human], as in (3b).

(3) a. La gata que buscó a la niña entró inesperadamente. [+animate, +human]
   ‘The cat that looked for DOM the girl entered unexpectedly’
   b. La niña que buscó (a) la gata entró inesperadamente. [+animate, −human]
   ‘The girl that looked for (DOM) the cat entered unexpectedly’
In object RCs (OVS), with the reduced pronoun ‘que’ (‘that’), which is the most frequent in Peninsular Spanish, the subject is postponed. Therefore, the postverbal constituent (the subject) does not have DOM, regardless of the [+human] or [−human] feature, as (4) shows.

(4) Object relative clauses (OVS)  
a. La gata que buscó la niña entró inesperadamente. [+animate, +human]  
   The cat that looked for the girl entered unexpectedly  
   ‘The cat that the girl looked for entered unexpectedly.’

b. La niña que buscó la gata entró inesperadamente. [+animate, −human]  
   The girl that looked for the cat entered unexpectedly  
   ‘The girl that [the cat looked for/looked for the cat] entered unexpectedly.’

Notably, the RCs in (3a) and (4a) are structurally ambiguous, but only temporarily until the onset of the postverbal constituent (crucially, the presence/absence of DOM) signals its syntactic function. However, the RC in (4b) is permanently ambiguous since the [−human] postverbal NP within the RC can be interpreted in two ways, as follows: (i) as a postponed subject; or (ii) as an object, given that it can be unmarked (optional marking). For these reasons, the example in (3b) would only be unambiguous without the DOM marker.

Regarding Catalan, the scope of structural ambiguity in RCs is broader than in Spanish due to the restricted DOM system. DOM in standard Catalan (a morpheme, as in Spanish) has a much more limited distribution than in Spanish: only personal pronouns are always marked, as shown in the contrast of (5) (Institut d’Estudis Catalans 2016). This restriction fits within the Definiteness Scale (Aissen 2003) in (1b), which proposes that the more definite and specific an object is, the more likely it is to be marked.

(5)  
a. La va veure a ella. [+human, +personal pronoun]  
   CL AUX saw DOM she  
   ‘He/she saw her.’

b. La Carla va veure la noia. [+animate, +human]  
   The Carla AUX seen the girl  
   ‘Carla saw the girl.’

Given the general absence of DOM in standard Catalan, whether it is a [+animate, +human] or [+animate, −human] direct object, the structural ambiguity described for RCs in Spanish is extended to other contexts. In Catalan, subject RCs with a [+human] object, such as (6a), coincide with object RCs with [+animate, −human] objects, such as (7a), and subject RCs with a [+animate, −human] object, such as (6b), coincide with object RCs with [+animate, +human] objects such as (7b). In other words, postverbal subjects may be interpreted as subjects or objects when subject–verb agreement does not disambiguate the sentence.

(6) Subject relative clauses (SVO)  
a. La gata que va buscar la nena va entrar inesperadament. [+animate, +human]  
   The cat that looked for the girl entered unexpectedly  
   ‘The cat that looked for the girl entered unexpectedly.’

b. La nena que va buscar la gata va entrar inesperadament. [+animate, −human]  
   The girl that looked for the cat entered unexpectedly  
   ‘The girl that looked for the cat entered unexpectedly.’

(7) Object relative clauses (OVS)  
a. La gata que va buscar la nena va entrar inesperadament. [+animate, −human]  
   The cat that looked for the girl entered unexpectedly  
   ‘The cat that the girl looked for entered unexpectedly.’

b. La nena que va buscar la gata va entrar inesperadament. [+animate, +human]  
   The girl that looked for the cat entered unexpectedly  
   ‘The girl that the cat looked for entered unexpectedly.’

Importantly, the non-standard Catalan DOM system displays alternation allowing [+animate] objects to be optionally case marked. In this scenario, the ‘a’ marker would appear in (6a) preceding the object ‘la nena’ and in (6b) preceding ‘el gat’, resulting in the non-standard Catalan sentences: La gata que va buscar a la nena... and La nena que va buscar...
... This is the variety that many Catalan speakers usually employ in everyday exchanges and conversations, and there is a debate as to whether DOM optionality is a result of prolonged language contact with Spanish, an internal linguistic process, or a convergence of both processes (Benito 2023; Perpiñán 2018). Briefly, literate speakers are bidialectal and know and use, depending on the register, both varieties of Catalan (the standard variety with virtually no DOM and the non-standard variety with optional DOM). Therefore, it is crucial to consider this idiosyncrasy in order to better understand the outcomes in bilingual Spanish.

Many studies on CLI between Spanish and English show that DOM is vulnerable to CLI in non-native and bilingual speakers (see Hopp and Arriaga 2016; Jegerski 2015; Montrul et al. 2015). The fact that DOM in Spanish is optional with [+animate, −human] objects giving rise to ambiguity, added to the fact that standard Catalan is much more restrictive regarding DOM, makes it an ideal domain to test how early Spanish and Catalan–Spanish balanced bilinguals interpret and process these potentially ambiguous sentences.

1.2. Online Processing of DOM and RCs in Native and Bilingual Speakers

1.2.1. Differential Object Marking (DOM)

There are a number of studies examining DOM in L2 Spanish (Guijarro-Fuentes 2012; Hopp and Arriaga 2016; Jegerski 2015; Montrul 2004; Sagarra et al. 2020) and heritage Spanish in English contexts (Montrul 2004; Montrul and Bowles 2009; Sagarra et al. 2017; Thane 2024a) illustrating that DOM is indeed a challenging phenomenon to acquire, and susceptible to CLI marked by inherent variability. However, to the best of our knowledge, only three studies have focused on Catalan–Spanish bilingual speakers.

Puig-Mayenco et al. (2018) carried out a bidirectional study testing the animacy property through a self-paced reading task. In Catalan, the Catalan-dominant group did not display sensitivity to the violation of the [+DOM] condition in human objects, which consisted of proper names preceded by the ‘a’ DOM (ungrammatical in standard Catalan, as standard Catalan only permits DOM with personal pronoun objects). The Spanish-dominant group did show sensitivity to the [−DOM] condition in human objects (which is grammatical in the tested language, Catalan, but ungrammatical in their dominant language, Spanish). These findings indicate CLI from Spanish onto Catalan that, in addition, varies depending on language dominance. In Spanish, Catalan-dominant speakers are sensitive to the [+DOM] condition (grammatical in Spanish, but ungrammatical in their dominant language) and Spanish-dominant speakers are sensitive to the [−DOM] condition (ungrammatical in Spanish). Thus, in Spanish, CLI is only apparent for DOM in Catalan-dominant bilinguals.

Bidirectional CLI in Catalan and Spanish has also been demonstrated with offline data. In an acceptability judgment task, Puig-Mayenco et al. (2018) attested optionality in Catalan, as participants accepted [+DOM] and [−DOM] sentences and found an over-acceptance of [−DOM] in Spanish. Similarly, Perpiñán (2018) argues that there is optionality when accepting DOM in Catalan and certain CLI when producing it, mostly (but not only) in Spanish-dominant bilinguals. Contrary to these results, Guijarro-Fuentes (2012) production study in Spanish revealed that Catalan–Spanish bilinguals overuse ‘a’ marking in optional contexts, which, according to them, does not fit the CLI hypothesis because Catalan has a more restricted DOM system. However, this could be due to the environment, as their participants lived in Barcelona, an area with more contact with Spanish than other regions. It would thus be in line with the studies above.

In summary, there is no consensus regarding the role of CLI between the two DOM systems among bilinguals. Regarding its effect on online processing, asymmetric results emerge depending on the target language and language dominance.
1.2.2. Relative Clauses (RCs)

RCs are an excellent context in which to study the role of case-marking because the flexible word order allowed in some languages produces potentially ambiguous sentences that can crucially be disambiguated with case-marking. As described above, in Spanish, DOM disambiguates RCs in which subject and object coincide in person and number. When DOM is optional, i.e., with animal objects, the RC remains ambiguous. In standard Catalan, which does not have overt case-marking (particularly in the cases under scrutiny), all RCs become ambiguous unless subject-verb agreement elucidates the meaning.

Studies point out that word order, together with case-marking, seems more crucial in some languages than others. On one hand, Hopp (2006) studied word order in German L2 learners and found evidence of online sensitivity to case-marking in RCs. Importantly, the native-like group showed a preference for subject-first constructions. On the other hand, Mitsugi and MacWhinney (2015) noticed that L2 learners of Japanese obviated attention to word order and relied on case markers to assign verb argument roles.

Research also indicates that subject RCs are easier to process than object RCs (King and Just 1991 for English; Holmes and O'Regan 1981 for French; Frazier 1987 for Dutch, among many others), and easier to acquire (Izumi 2003 in L2 English; Hawkins 1989 in L2 French; Liceras 1986 in L2 Spanish). Moreover, RCs have been shown to be influenced by the animacy of the subject and the object. Through eye-movement-monitoring experiments (Mak et al. 2006) and a self-paced word-by-word reading task (Traxler et al. 2002), the authors observed that when both the subject and the object are animate, or when the preverbal NP is animate and the postverbal NP is inanimate, the difficulty in interpreting object relative clauses increases. The difficulty decreases when the preverbal NP is inanimate and the postverbal NP is animate, or when both are inanimate (Mak et al. 2006). The impact of word order and animacy in sentence processing has also been found in matrix sentences. Casado et al. (2005), using ERPs, found that inanimate NPs in the first position posed an initial difficulty in interpreting them as subjects. Similar results are demonstrated in a study by Yoza (2016), who replicated the study of Casado et al. (2005) using a self-paced reading task.

Although there is extensive research about RC processing in languages such as English, there is little research on Spanish and even less on Catalan. Using eyetracking while reading, Betancort (2006) investigated RC processing in Spanish by manipulating the order and animacy of NP antecedents (“atleta” vs. “enfermedad”) to test if the animacy semantic feature influences the assignation of theta-roles in Spanish (“Conocí [al atleta/la enfermedad] que venció finalmente [al corredor/el corredor] el año pasado”, ‘They knew {the athlete/the illness} {that beat the runner/that the runner beat} finally last year’). Remember that an animate object would trigger DOM, which would be the clue for interpretation. The results showed that OVS sentences are more cognitively taxing to process than SVO sentences, which is consistent with prior work, and suggest that, although animacy is not initially used to process RCs, it is used in the subsequent reanalysis (from the default initial SVO analysis to the target OVS). Later, Betancort et al. (2009) claimed that SVO subject RCs are not only easier to process but also more preferred than OVS object RCs, and that this choice is modulated by animacy. Similar to the results of previous research, having two animate NPs was more cognitively taxing than having a preverbal inanimate object and a postverbal animate object.

Del Río et al. (2012) studied whether increased processing costs in Spanish object RCs are due to working memory demands for the establishment of long-distance dependencies or to difficulties processing non-canonical structures. The findings suggest that a preference for canonical structures influences interpretation and, congruent with previous research, animacy affects sentence processing.

As for the role of DOM, Perpiñán and Moreno Villamar (2013) investigated word order and DOM in RC interpretation in Spanish with natives and heritage speakers, concluding that natives prefer word order over DOM. Similar results are shown in Sagarra et al. (2017), who found that native and heritage speakers of Spanish are less accurate in interpreting
object RCs (which do not display DOM) than subject RCs (which overtly display DOM). However, processing data showed that they were slower in subject RCs due to the presence of DOM.

Regarding Catalan, its pattern resembles data from other languages. Gavarró et al. (2012) investigated RCs in children and found differences between subject and object RCs, both in comprehension and production. However, we are not aware of any study about the role of animacy in this language. This gap is not surprising, as there is no overt case-marking in standard Catalan. Consequently, subjects and objects are formally the same, whatever animacy they encode.

Given the differences between standard Catalan and Spanish regarding animacy as a trigger of obligatory or optional DOM, particularly relevant within RCs, investigating Spanish monolinguals and early Catalan–Spanish-balanced bilinguals will allow us to test if they resort to different strategies to solve ambiguity (word order and/or DOM). It will also shed light on the bilinguals’ processing strategies because bilingualism places an increased burden on memory processes (Marian and Fausey 2006; Sorace 2016) and may be more visible when processing complex structures. Regarding the role of CLI in bilingualism, it is worth highlighting that it also affects early bilinguals. As mentioned, previous work with bilinguals of the same origin as our participants (from a bilingual community like Catalonia, where bilingual practices and language contact constantly occur) found selective CLI, even among balanced bilinguals, that in the case of DOM ended in an optional outcome (Perpiñá 2018; Puig-Mayenco et al. 2018). At the individual level, optionality has been proposed to characterize bilingual grammars (Sorace 2016); at the social level, bilingual settings enhance alterations of grammatical knowledge and diversity (Meisel 2021). Moreover, and importantly, given that early Catalan–Spanish bilinguals also know and use the non-standard Catalan variety, which displays DOM fluctuation, there is more overlapping between languages that could further impact the Spanish DOM system emphasizing optionality. Finally, it has been alleged that balanced bilinguals are more prone than unbalanced bilinguals to show coactivation of grammatical representations (Putnam and Sánchez 2013) and this would favor the emergence of CLI and result in a certain level of optionality. More recently, Putnam et al. (2018) have explored the connection between language dominance and crosslinguistic influence (CLI), suggesting that language dominance plays a crucial role in determining the extent and direction of CLI between languages. They propose that CLI is more likely to occur from the dominant language to the non-dominant language. Balanced bilingual speakers, on the other hand, may experience bidirectional CLI effects. While this theory is interesting, it is not directly applicable to our study design, as we focus solely on Spanish and a population composed exclusively of balanced bilingual individuals.

1.3. The Study: Aims and Predictions

Past research has examined the processing and interpretation of relative clauses and DOM separately. In addition, little research has addressed how bilinguals deal with either of these two topics, particularly early bilinguals of the nature we focus on, who are highly balanced and raised in a bilingual community. The current study is unique in exploring bilingual processing and the interaction between two phenomena. We test a group of highly proficient Catalan–Spanish bilinguals to examine whether DOM realization and word order in relative clauses in Spanish affect processing and interpretation. Concerning DOM, we concentrate on the top portion of the animacy scale concerning the human/non-human distinction. This is where we can examine DOM optionality.

An additional innovative aspect of the current study addresses how DOM is processed within RCs and constitutes a challenging scenario since the interpretation of the RC’s object relies solely on the presence/absence of DOM. Indeed, the study of DOM in more complex structures in comparison to the ones previously addressed in the literature, that is, RCs, offers a privileged situation to test DOM since, different to a simple sentence, in an RC, it is the very presence (or absence) of the ‘a’ DOM marker in the postverbal argument that
signals that the argument is a subject or an object. In a Spanish simple sentence like “La
tenista persigue a la periodista” (‘The tennis player chases the journalist’) the reader can
anticipate, without error, that the postverbal element will be, and can only be, an object
(before reading the ‘a’ marker); in addition, in an OVS simple sentence, again, there is
no room for error in anticipating the syntactic function of the postverbal argument (i.e., a
subject): “A la periodista la persigue la tenista” (recall that this sentence involves a typical
object topicalization resulting in clitic doubling, which is not relevant for our argumentation
here). Hence, the interpretive/functional distinction in Spanish transitive simple sentences
does not only rely only on DOM but also on the restrictions imposed by sequencing.

Specifically, we formulate the following research questions and predictions:

1. RQ1: What are the patterns of interpretation and processing of obligatory versus
   optional DOM in Spanish among monolingually raised Spanish speakers? What are
   the patterns observed in the case of Catalan–Spanish bilinguals?

   For DOM, we expect that monolingual’s processing and interpretation of Spanish
human objects (obligatory marked) and animal objects (optional marked) will be driven by
the overt marking realization; we predict that obligatory overt marking, which can be only
satisfied one way, will be easier to process (shorter reaction times) than optional marking.
In the case of bilinguals, this pattern will not emerge due to the CLI effect resulting from the
absence of DOM with both human and non-human objects in the other language, Catalan.
Furthermore, the presence of a non-standard variety, featuring uneven or optional DOM,
will intensify and clarify the CLI effect.

2. RQ2: Are subject RCs easier to interpret and process than object RCs in Spanish? If
   so, is this ease of processing influenced by the presence or absence of DOM when
   interpreting and parsing relative clauses, both in monolinguals and bilinguals?

   For RCs, we anticipate greater difficulty in processing and interpreting OVS object RCs
compared to SVO subject RCs both in monolingual and in bilingual speakers. In the case
of Catalan–Spanish bilinguals, differing patterns from those of Spanish monolinguals are
expected due to the interaction with DOM. The absence or unevenness of DOM in the other
language may mitigate differences in reading times between both types of relative clauses.
Finally, given the complexity of the structural phenomenon under discussion, it
is conceivable that this complexity contributes to cognitive demands that may impact
bilingual speakers (see Sorace 2016, regarding cognitive taxing among bilinguals), thus
leading to optionality as a characteristic feature of bilingual grammars.

2. Materials and Methods

2.1. Participants

Thirty-two bilingual speakers of Catalan and Spanish, along with 32 individuals
raised monolingually in Spanish (mean age = 22.7 years), voluntarily participated in the
experiment. Spanish speakers were university students born and raised in a monolingual
province of central Spain (mean age = 23.5 yrs). Participants with a first language other
than Spanish were excluded. As university students, our group of individuals raised mono-
lingually also had a knowledge of English, which they had studied at various school levels
as a foreign language. According to the LEAP-Q questionnaire administered (Marian et al.
2007), their proficiency in English ranged from low-intermediate to upper-advanced levels.

Bilingual Catalan–Spanish speakers were also from university contexts, which means
that they had mastered the standard Catalan variety. However, they had also been exposed
to a non-standard Catalan that uses DOM, or at least accepts it (Bel and García-Alcaraz
2018; Perpiñán 2018; Puig-Mayenco et al. 2018), since it is a variety that is spreading
(Benito 2023).1 They completed a comprehensive bilingual profile questionnaire specifically
tailored for our bilingual community within our research group that was based on previous
questionnaires constructed within our sociolinguistic context. The questionnaire covered
different levels of demographic and sociolinguistic information and usage such as the
frequency of language use in daily life across different contexts (family, friends, language
practices in various environments, primary and secondary school, where specific practices are derived from language educational policies of the Autonomous Government, etc.; see Appendix B). Participants were born and raised in Catalonia, and the majority belonged to bilingual families: 41.6% acknowledged a bilingual home situation, with 38.9% declaring Catalan as their home language and 19.4% declaring Spanish. In the three cases, only participants with excellent knowledge and broad use of both languages were included. In any case, all participants had continuous and extensive exposure to both languages before the preschool level. They identified themselves as balanced bilinguals and admitted to feeling as comfortable in Spanish as in Catalan. It should be noted that the average daily usage of Catalan was slightly higher than Spanish in childhood, puberty, and adulthood (Table 1). It is also worth remembering that our participants were university students, and that Catalan is the primary language of instruction at different educational levels (for the significance of instruction on the acquisition of DOM in bilinguals, particularly heritage speakers of Spanish, see Thane 2024b).

Table 1. Bilinguals’ self-reported sociolinguistic information.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onset age of exposure to oral Catalan (mean in years)</td>
<td>0.45</td>
<td>0.99</td>
</tr>
<tr>
<td>Onset age of exposure to oral Spanish (mean in years)</td>
<td>2.89</td>
<td>2.10</td>
</tr>
<tr>
<td>Self-reported Catalan level (out of 4)</td>
<td>3.99</td>
<td>0.06</td>
</tr>
<tr>
<td>Self-reported Spanish level (out of 4)</td>
<td>3.98</td>
<td>0.06</td>
</tr>
<tr>
<td>Average daily use of Catalan in childhood (out of 6) *</td>
<td>4.57</td>
<td>0.99</td>
</tr>
<tr>
<td>Average daily use of Catalan in puberty (out of 6) *</td>
<td>4.44</td>
<td>1.02</td>
</tr>
<tr>
<td>Average daily use of Catalan in adulthood (out of 6) *</td>
<td>4.35</td>
<td>0.95</td>
</tr>
</tbody>
</table>

*From 1 to 6 (1 = only Spanish; 6 = only Catalan).

Both groups were similar in terms of education (university students), handedness (right) and age (19–25 years old). In addition, given the type of online task administered (a self-paced reading), we controlled between-group homogeneity in reading times (because some participants are faster readers than others) and working memory (WM) (because it can intervene in morphosyntactic processing). To check the homogeneity in sentence reading times, independent samples t-tests on reading speed scores were calculated and homogeneity was attested ($t(60) = 1.37, p = 0.379$). To assess WM, we adapted the Wechsler Adult Intelligence Scale’s (WAIS) letter–number sequencing test. The results revealed no significant differences between the monolingual and the bilingual group (independent samples t-test $t(60) = 1.89, p = 0.639$).

2.2. Stimuli and Procedure

To assess the processing of DOM in subject and object RCs, a self-paced reading task with two conditions, the order of the RC (SVO/OVS) and the type of animate object (human/animal), was designed. The following experimental sentences exemplify each level condition: (8) illustrates subject RCs (SVO); (9) exemplifies object RCs (OVS) (as explained in a previous section, (9b) is ambiguous concerning the subject and an object RC):

(8) Subject RC (SVO)
   a. La gorila que abrazó a la bióloga sonreía incesantemente. (animal S, human O)
   ‘The gorilla that hugged the biologist smiled unceasingly’
   b. La bióloga que abrazó a la gorila sonreía incesantemente. (human S, animal O)
   ‘The biologist that hugged the gorilla smiled unceasingly’

(9) Object RC (OVS)
   a. La gorila que abrazó la bióloga sonreía incesantemente. (animal O, human S)
   ‘The gorilla that the biologist hugged smiled unceasingly’
   b. La bióloga que abrazó la gorila sonreía incesantemente. (human O, animal S)
   ‘The biologist that the gorilla hugged smiled unceasingly’
The task was a self-paced reading task programmed in the E-Prime software (version 1.4.1.1). It consisted of a word-by-word, non-cumulative, moving window presentation (participants saw each word replaced by underscores) of 48 experimental sentences (12 per RC order), 84 fillers, and three practice items. The filler and experimental sentences were grouped and randomized in twelve blocks (Latin square design), and participants were assigned to a different list so that each participant read each item in only one condition. The fillers included sentences with intrasentential subject pronouns and subject–verb agreement. Given the focus of the current work, we included feminine singular and plural NPs to mark DOM saliency with an unbound ‘a’ marker (‘a la’, ‘a las’); specifically, we avoided the less-salient masculine singular contraction ‘al’ (i.e., the contraction of ‘a + el’).

The experimental sentences were fully lexically matched to ensure comparability. The nouns were 2–4 syllables long and appeared once, and the verbs (transitive RC verbs, intransitive matrix verbs) were 2–4 syllables long and appeared twice, but never in the same combination of words. Importantly, the RC verbs were thematically reversible for interpreting humans and animals as subjects or objects (based on ratings from six Spanish monolinguals). This reversibility is evident in examples (8)–(9) above. Finally, the lexical selection of the nouns and verbs was based on word length, semantics (reversibility), and familiarity (according to three Spanish textbooks and a pilot vocabulary test given to similar participants).

All experimental sentences were grammatical, and all objects within SVO subject RCs were DOM marked regardless of the [+human] feature (see examples in (8) above). If ungrammatical DOM objects had been included, this would have created an asymmetry, as the lack of DOM with [+human] objects leads to ungrammaticality. On the contrary, the lack of DOM with [−human] objects not only gives rise to a grammatical outcome, since marking is optional, but it also generates an ambiguous sentence because it can be interpreted as an object RC with a postverbal subject. However, the condition OVS with a human object, as in (9b), was inevitably ambiguous. This ambiguous condition was not considered when comparing word order (SVO vs. OVS), because our participants’ interpretations can be variable. Therefore, only subject/object RCs with human postverbal nouns were compared, and, consequently, only pairwise comparisons within this condition were considered.

All sentences were followed by a post-stimulus verification task in the form of a yes–no question. Half of the questions assessed comprehension of the sentence’s content and attention to the experiment through the accuracy of sentential questions (e.g., ¿La gorila sonreía? ‘Did the gorilla smile?’). The other half involved questions assessing RC interpretation (e.g., ¿La gorila abrazó a la bióloga? ‘Did the gorilla hug the biologist?’).

Therefore, two critical measures were obtained from the task, as follows: offline interpretation regarding RC questions; and online reading times (RTs) at different word regions.

Online processing was evaluated with reading time (RT) data for each word region. However, analyses were run from the last word of the relative clause, i.e., the noun head of the object NP (in SVO subject RCs items) or the postponed subject NP (in OVS object RCs items); analyses on the subsequent matrix verb and the following adverb are also reported. We note that reading times in the sentence-final region, i.e., the adverb, should be cautiously interpreted since they reflect wrap-up effects (Just and Carpenter 1980). Reading times faster than 200 ms or slower than 2500 ms were excluded according to general practices in self-paced reading studies. If participants displayed higher RTs in a given region in a particular condition, this would be a symptom of interpretation difficulties on that condition.

3. Results

Mixed-effects ANOVAs accompanied by pairwise contrasts were performed using SPSS software (IBM SPSS Statistics 22, version 22.0), one per each dependent variable: interpretation of RCs questions and RTs at the critical region (the last N of the RC), at the possible spill-over region (the main VERB, one word after the critical noun is encountered).
and at the final wrap-up region (the ADVERB; two words after the critical noun). The types of RC and DOM, and all their possible interactions, were included as fixed factors, and Subject and Item were included as random factors. The first mixed-effects ANOVA followed a binomial distribution (logit link) and the other three followed a gamma distribution (log link). Alpha level was 0.05 and pairwise comparisons were calculated with a Bonferroni correction. Given our goal, inferential analyses were conducted separately for each group (monolinguals and bilinguals) to ascertain whether different processing behavior patterns arose within each group’s linguistic profile. Thus, no overt comparisons between groups are offered.  

3.1. Content Comprehension and RC Interpretation Questions

Content comprehension questions were only used to assess participants’ attention to the experiment. Response accuracy was 87.8% (SD = 6.5) for the monolingual group and 90.63% (SD = 3.3) for the bilingual group, both indicating high levels of accuracy. These results confirmed that participants indeed paid attention, so no participant was excluded. RC questions were used to assess the RC interpretation, depending on word order and DOM. Descriptive statistics can be found in Table 2 and are graphically displayed in Figure 1. It is worth remembering that a postverbal animal subject can be interpreted as a true object, given the optional marking that this kind of noun displays.

Table 2. Descriptive statistics for RC interpretation questions in percentages (the percentage corresponds to the underlined subject/object).

<table>
<thead>
<tr>
<th>Subject RC (SVO)</th>
<th>Object RC (OVS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Human Postverbal Noun (Object)</strong></td>
<td><strong>Animal Postverbal Noun (Object)</strong></td>
</tr>
<tr>
<td>Obligatory DOM (e.g., La gorila que abrazó a la bióloga...)</td>
<td>Optional DOM (e.g., La bióloga que abrazó a la gorila...)</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>---</td>
<td>----</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>87.5</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>93.75</td>
</tr>
</tbody>
</table>

Figure 1. RC question interpretation results for monolinguals and bilinguals (in % in the y-axis). The human feature refers to the postverbal noun.

The output of the analyses for RC questions is provided in Table 3. The significant main effect of Type of RC showed that object RC questions were interpreted more heterogeneously than subject RC questions, as shown by the descriptive results in Table 2, both for monolingual and bilingual speakers (remember that half of the object RCs were
ambiguous). A general main effect of DOM only emerged among bilinguals. Interestingly, the interaction of Type of RC × DOM only appeared in the monolingual group, showing a selective sensitivity to DOM. The pairwise comparisons of this interaction revealed the two following patterns. First, from the perspective of Type of RC, subject RCs with human postverbal nouns (objects) are better interpreted than object RCs with human postverbal nouns (subjects) ($t(136) = 0.20, p = 0.025$); on the other hand, subject and object RCs with postverbal animal nouns also showed differences in interpretation ($t(136) = 0.79, p = 0.001$), but this was due to the inherent ambiguity of object RCs with postverbal animal nouns. Second, from the perspective of DOM and only within object RCs (OVS sequences), differences between the interpretation of human and animal postverbal elements occurred ($t(136) = 0.51, p = 0.001$), given that the latter (the ambiguous non-marked postverbal animal elements) are mainly interpreted as objects, which is a possible outcome. These results are graphically displayed in Figure 1.

Table 3. Mixed-effects ANOVAs for RC interpretation question results by group.

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monolinguals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of RC</td>
<td>42.618</td>
<td>1</td>
<td>136</td>
<td>&lt;0.001</td>
<td>0.742</td>
</tr>
<tr>
<td>DOM</td>
<td>1.232</td>
<td>1</td>
<td>136</td>
<td>0.269</td>
<td>0.020</td>
</tr>
<tr>
<td>Type RC × DOM</td>
<td>15.022</td>
<td>1</td>
<td>136</td>
<td>&lt;0.001</td>
<td>0.207</td>
</tr>
<tr>
<td><strong>Bilinguals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of RC</td>
<td>46.504</td>
<td>1</td>
<td>136</td>
<td>&lt;0.001</td>
<td>0.820</td>
</tr>
<tr>
<td>DOM</td>
<td>8.115</td>
<td>1</td>
<td>136</td>
<td>0.050</td>
<td>0.140</td>
</tr>
<tr>
<td>Type of RC × DOM</td>
<td>3.477</td>
<td>1</td>
<td>136</td>
<td>0.064</td>
<td>0.060</td>
</tr>
</tbody>
</table>

3.2. Reading Times

The descriptive statistics for the monolingual Spanish speakers’ RT scores can be found in Table 4. The three mixed-effects ANOVAs on RTs in the different regions revealed several significant effects, mainly at the last N of the RC, the critical region, with distinct patterns between groups. Starting with the monolingual group (see Table 4), in this critical region, there was a significant main effect of Type of RC (participants were slower at processing object RCs than subject RCs, $p = 0.012$) and of DOM (participants were slower at processing human postverbal nouns than non-human postverbal nouns, $p = 0.001$); the latter outcome might be due to the postverbal human subjects in OVS sentences that participants sometimes interpreted as objects (73.7% interpret it as a subject; see Table 2). No interaction was revealed. No delayed effects appeared in the following two segments (matrix verb and final adverb) within this group (see Table 5). The results are also graphically illustrated in Figures A1 and A2 (see Appendix A).

Table 4. Descriptive statistics of RTs (means in milliseconds) in different regions: monolingual speakers.

<table>
<thead>
<tr>
<th>Region</th>
<th>Human Postverbal Noun (Object)</th>
<th>Animal Postverbal Noun (Object)</th>
<th>Human Postverbal Noun (Subject)</th>
<th>Animal Postverbal Noun (Subject/Object)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obligatory DOM</td>
<td>Optional DOM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(e.g., La gorila que abrazó a la bióloga...)</td>
<td>(e.g., La bióloga que abrazó a la gorila...)</td>
<td>(e.g., La bióloga que abrazó a la bióloga...)</td>
<td>(e.g., La bióloga que abrazó a la gorila...)</td>
</tr>
<tr>
<td>Last N RC</td>
<td>M 524</td>
<td>M 566</td>
<td>M 735</td>
<td>M 595</td>
</tr>
<tr>
<td>'bióloga/gorila'</td>
<td>SD 289</td>
<td>SD 523</td>
<td>SD 694</td>
<td>SD 653</td>
</tr>
<tr>
<td>Verb</td>
<td>M 555</td>
<td>M 517</td>
<td>M 657</td>
<td>M 529</td>
</tr>
<tr>
<td>'sonreía'</td>
<td>SD 346</td>
<td>SD 263</td>
<td>SD 401</td>
<td>SD 241</td>
</tr>
<tr>
<td>Adverb</td>
<td>M 851</td>
<td>M 969</td>
<td>M 873</td>
<td>M 833</td>
</tr>
<tr>
<td>‘incesantemente’</td>
<td>SD 615</td>
<td>SD 640</td>
<td>SD 534</td>
<td>SD 608</td>
</tr>
</tbody>
</table>
Table 5. Mixed-effects ANOVAs for RT by region: monolingual speakers.

<table>
<thead>
<tr>
<th>Region</th>
<th>Type of RC</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last N RC</td>
<td>DOM</td>
<td>6.508</td>
<td>1</td>
<td>283</td>
<td>0.011</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>14.602</td>
<td>1</td>
<td>283</td>
<td>&lt;0.001</td>
<td>0.563</td>
</tr>
<tr>
<td></td>
<td>Type of RC</td>
<td>1.868</td>
<td>1</td>
<td>283</td>
<td>0.173</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>DOM</td>
<td>3.154</td>
<td>1</td>
<td>278</td>
<td>0.077</td>
<td>0.498</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>0.356</td>
<td>1</td>
<td>278</td>
<td>0.551</td>
<td>0.056</td>
</tr>
<tr>
<td>Verb</td>
<td>DOM</td>
<td>3.161</td>
<td>1</td>
<td>278</td>
<td>0.077</td>
<td>0.499</td>
</tr>
<tr>
<td></td>
<td>Type of RC</td>
<td>0.079</td>
<td>1</td>
<td>277</td>
<td>0.779</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>1.670</td>
<td>1</td>
<td>277</td>
<td>0.197</td>
<td>0.256</td>
</tr>
</tbody>
</table>

For the Catalan–Spanish bilingual group (descriptive statistics of the RT scores are illustrated in Table 6) no main effects were observed in the critical region of interest (the last noun of the RC) (see ANOVA data in Table 7). Still, an interesting significant interaction of the Type of RC × DOM revealed that, from the perspective of RC, the participants were slower at processing human postverbal nouns ($t(285) = -51.61, p = 0.015$) in OVS sequences in comparison with SVO ones. This result indicates the following: (1) within the human condition (postverbal noun), which undoubtedly distinguishes between subject and object syntactic functions due to the obligatoriness of case-marking, the presence/absence of the ‘a’ marker plays a crucial role for bilinguals in identifying the object; and (2) within this same condition, which allows for an overt comparison between the two types of RCs, the object RCs are processed more slowly than subject RCs. Within the animal noun condition (optional marking), the pairwise comparisons showed also that subject RCs were interpreted more quickly than object RCs ($t(285) = 70.99, p = 0.006$). However, this result is misleading and cannot be used to compare types of RCs because the object RCs were ambiguous. From the perspective of DOM, the pairwise comparisons revealed that, in subject RCs, optional DOM (animal objects) is processed more quickly than obligatory DOM (human objects) ($t(285) = -89.90, p = 0.001$), in the opposite manner to the pattern observed among monolinguals. In other words, it appears that bilinguals are sensitive to DOM as a cue for recognizing human objects (regardless of the non-DOM nature of their L1 Catalan, at least as far as the standard variety is concerned). On the other hand, the presence of the ‘a’ marker in an optional context does not interfere with the identification of a non-human object. This pattern of sensitiveness will be further developed in the next section.

Table 6. Descriptive statistics of RTs (means in milliseconds) in different regions: bilingual speakers.

<table>
<thead>
<tr>
<th>Region</th>
<th>Human Postverbal Noun (Object)</th>
<th>Animal Postverbal Noun (Subject)</th>
<th>Subject RC (SVO)</th>
<th>Object RC (OVS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obligatory DOM (e.g., La gorila que abrazó a la bióloga...)</td>
<td>Optional DOM (e.g., La bióloga que abrazó a la gorila...)</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Last N RC</td>
<td>538</td>
<td>373</td>
<td>500</td>
<td>255</td>
</tr>
<tr>
<td>‘bióloga/gorila’</td>
<td>527</td>
<td>297</td>
<td>485</td>
<td>231</td>
</tr>
<tr>
<td>‘incensamente’</td>
<td>812</td>
<td>696</td>
<td>856</td>
<td>705</td>
</tr>
</tbody>
</table>
Table 7. Mixed-effects ANOVAs for RT by region: bilingual speakers.

<table>
<thead>
<tr>
<th>Region</th>
<th>Type of RC</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last N RC</td>
<td>DOM</td>
<td>0.213</td>
<td>1</td>
<td>285</td>
<td>0.645</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>14.458</td>
<td>1</td>
<td>285</td>
<td>&lt;0.001</td>
<td>0.674</td>
</tr>
<tr>
<td>Verb</td>
<td>DOM</td>
<td>5.443</td>
<td>1</td>
<td>282</td>
<td>0.020</td>
<td>0.445</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>0.024</td>
<td>1</td>
<td>282</td>
<td>0.877</td>
<td>0.001</td>
</tr>
<tr>
<td>Adverb</td>
<td>DOM</td>
<td>0.277</td>
<td>1</td>
<td>283</td>
<td>0.599</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>Type of RC × DOM</td>
<td>1.740</td>
<td>1</td>
<td>283</td>
<td>0.188</td>
<td>0.229</td>
</tr>
</tbody>
</table>

Concerning the two mixed-effects ANOVAs on the subsequent regions, the following effects emerged: (1) a sole main effect of DOM at the verb region, such that bilinguals showed spill-over effects evinced by the slower processing of RCs with human than with non-human postverbal nouns ($t(282) = 42.19, p = 0.022$); and (2) a delayed significant effect at the adverb closing region, pointing to the well-known contrast between the processing cost of the two types of RCs—subject RCs were processed more quickly than object RCs ($t(283) = −95.89, p = 0.031$).

4. Discussion and Conclusions

This study examined the impact of DOM on the processing distinctions between subject and object RCs in Spanish monolingual and bilingual speakers (in contact with Catalan). We tested the hypothesis concerning the interaction of DOM and RC relative word order by manipulating two types of objects (+/-animate) representing obligatory and optional DOM marking, and two types of RCs, SVO and OVS. Unlike previous studies that exclusively compare [+/-animate] objects, we focused on the higher end of the animacy scale in this study, particularly on the optionality in DOM marking that distinguishes [+/-human, +animate] objects in Spanish.

While it is challenging to separately address our research questions because DOM and word order in RCs interact, we have strived for clarity by first discussing our findings regarding DOM and subsequently addressing findings regarding RCs and their interaction with DOM.

RQ1. What are the patterns of interpretation and processing of obligatory versus optional DOM in Spanish among monolingually raised Spanish speakers and Catalan–Spanish bilingual speakers?

Before exploring the distinction between obligatory and optional marking, we aimed to ascertain whether our participants exhibit sensitivity to DOM. To accomplish this, we initially relied on processing data, as outlined in Section 3.2. The monolingual speakers solved the interpretation early in the time-course of processing since effects are only revealed at the opening critical region, i.e., the main effect of DOM at the noun closing the RC indicates that the presence/absence of DOM elucidates the subject/object dichotomy (in human postverbal nouns). For bilinguals, several notable observations emerged. First, as anticipated, they displayed delayed spill-over effects—i.e., they appeared to spend more time and greater cognitive resources to perform the appropriate analysis. Second, the interaction of type of RC and DOM at the critical region suggests that they were taking DOM into account when interpreting human objects, despite the non-DOM character of Catalan, at least where standard variety is concerned. Third, the uneven and irregular experience with DOM in their other language, Catalan, particularly in its non-standard variety that all our bilinguals also speak, enables bilinguals to easily accommodate the optional marking of non-human animal objects in Spanish. Finally, although we did not perform direct statistical comparisons between groups, the results not only show that bilinguals are not slower than monolinguals (see row data in Tables 4 and 6), but also
demonstrate that bilinguals exhibited faster RTs than monolinguals across most conditions and regions, except for subject RCs with postverbal human objects at the initial region and object RCs with postverbal human subjects at the final region.

Our prediction under RQ1 concerning contrasts between obligatorily marked human objects and optionally marked animal objects is supported by our findings. Specifically, monolingual processing shows overall slower rates with non-human postverbal nouns compared to human postverbal nouns in OVS sentences once they reach the last noun of the RC. Analysis of offline accuracy comprehension data also reveals that the presence of DOM aids in distinguishing between a [+human] postverbal subject or object. In instances where DOM is absent (i.e., postverbal [+/-human] subjects), participants tend to preferentially interpret them as objects, particularly when it is [-human]. This interpretation is plausible given their optional marking, as evidenced by over 80% of responses favoring this interpretation. The behavior observed with [-human] animal constituents suggests that optionality influences interpretation and imposes an additional processing cost (in SVO), and that DOM interacts with word order within the RC and drives interpretation. As for [+human] postverbal subjects, the object reading is banned, as they lack DOM and can only be interpreted as subjects. Nevertheless, a surprising 30% of our monolingual participants accepted the object interpretation despite the absence of DOM. We attribute this difference to the inherent complexity associated with object RCs (where subjects appear postverbally).

A second, complementary account for these findings is related to the semantic properties that intrinsically distinguish human and animal objects nouns and not the DOM marking associated with them (obligatory marking vs. optional marking, respectively). As reported for RCs including animate and inanimate subject and/or object nouns, animacy modulates the difficulty in processing, as stated in previous research (for RC see Betancort et al. 2009; Del Río et al. 2012; Mak et al. 2006; Traxler et al. 2002; for matrix sentences see Casado et al. 2005). It could be that ‘humanity’ mediates the subject–object interpretive division (i.e., in interpretive terms, it seems more likely to have a human subject and a non-human object than the opposite). The results from Figure 1 seem to support this explanation since animal postverbal objects (with DOM) are mainly interpreted as objects. In contrast, animal postverbal subjects (without DOM) are overwhelmingly interpreted as postverbal objects. Hence, postverbal animals in the NP2 position are consistently interpreted as postverbal objects far more frequently than as subjects, even though they display (optional) DOM. The two explanations, i.e., the role of optional DOM and the human/animal feature, cannot be sharply separated, but they conspire together inside this territory.

Regarding bilinguals, it appears that DOM optionality does not incur an additional cost. Combining comprehension and processing outcomes, bilinguals interpret human objects in a manner akin to monolinguals, showing that they have incorporated obligatory DOM into the Spanish case system. Second, unlike monolingual speakers, our bilinguals processed DOM-marked animal objects in subject RCs faster, showing that they can better adjust optional DOM, probably taking advantage of the uneven occurrence of DOM in non-standard Catalan. Third, like monolinguals, the bilingual group processed animal postverbal elements (without DOM) in object RCs faster than human postverbal elements. Thus, they seem to miss DOM with the latter, showing that, contrary to our prediction, bilinguals seem sensitive to DOM marking in Spanish, even though standard Catalan lacks DOM with both human and non-human objects. Our prediction was predicated on the assumption that a general absence of DOM in Catalan in the cases under scrutiny would result in no differences in the processing of the two types of nouns in Spanish due to CLI. However, the attested results suggest that the source of CLI does not stem from the standard variety but rather from the non-standard colloquial Catalan variety (see Benito 2023) spoken by our bilingual participants in everyday interactions and informal contexts. This variety has optional DOM with both types of object nouns, as discussed in Section 1.1 above. The optional DOM system in non-standard Catalan could lead to CLI in their Spanish DOM system, even though our bilinguals are highly proficient in both languages.
We lack precise information regarding the input and usage of the two varieties among our bilingual participants. Therefore, the statement underlined above should be approached cautiously, pending further data that adequately address and measure its impact. For now, we might argue that, if we are on the right path, our results do not show the vulnerability of Spanish DOM attested by Hopp and Arriaga (2016), Jegerski (2015) or Montrul et al. (2015), since the linguistic combination is not fully comparable (English and Spanish in their studies, Catalan and Spanish in ours). The few studies on Catalan DOM indicate that, due to various converging factors, such as extensive contact with Spanish, Catalan has evolved to include optional DOM with animate objects, which is extended depending on the relative dominance and exposure to languages (Benito 2023; Perpiñán 2018; Puig-Mayenco et al. 2018). The overlap between the Catalan and Spanish DOM systems is greater than was assumed from a prescriptive or ‘normative’ perspective within the Catalan sociolinguistic context. Finally, the interpretation of postverbal human subjects (without DOM) in object RCs (“La gorila que abrazó la biòloga...”) is free within the bilingual group. They are interpreted as true postverbal subjects or as illicit non-marked objects at the chance level. The latter option is possible if bilinguals are influenced by their non-DOM L1 (or by their optional L1 DOM system), suggesting that if CLI is amplified by coactivation, which has been posited as impacting balanced bilingual individuals showing bidirectional effects (Putnam and Sánchez 2013; Putnam et al. 2018), this influence is only discernible under specific DOM conditions.5 In short, from our findings, it can be inferred that CLI is not a uniform phenomenon that applies across the board but rather manifests overtly in some instances, covertly in others, and indirectly in still others (Perpiñán and Soto-Corominas 2021).

The Catalan (non-standard) DOM system shows alternation between DOM and non-DOM even at the highest levels of the animacy scale. This is a consequence of convergence between some language-internal processes and the effects of long-term contact with Spanish, a DOM language. In addition, Spanish is a well-established language in the Catalan community, so the pressure goes in two directions—from Spanish to Catalan and from Catalan to Spanish—potentially contributing to a specific variety of Spanish in this sociolinguistic non-homogeneous context. Moreover, if it is true that contact has a greater impact on the speaker’s non-dominant language (Silva-Corvalán 2014), it will be important to further study the DOM system with different profiles of bilingual speakers so that variable inputs can be managed effectively. This approach would also facilitate the testing of Putnam et al.’s (2018) proposal regarding the variable impact of CLI depending on language dominance. However, addressing this topic falls outside the scope of the current article.

RQ2. Are subject RCs easier to interpret and process than object RCs in Spanish? If so, is this ease of processing influenced by the presence or absence of DOM when interpreting and parsing relative clauses, both in monolinguals and bilinguals?

To address the initial aspect of this question concerning the relative complexity of object versus subject RCs, we exclusively consider the outcomes within the human nouns condition—specifically comparing SVO to OVS constructions, where the postverbal NP is either a human object or a human subject, (8a) and (9a), respectively. This approach enabled a comparison between both types of RCs. In this condition, processing results showed that object OVS RCs are read more slowly than subject SVO RCs, and this general pattern held for both monolingual and bilingual speakers. The evidence also indirectly emerged from offline comprehension data within the non-human condition (postverbal nouns): in object RCs, both groups interpreted the postverbal non-marked non-human—i.e., animal—mainly as objects (a possibility facilitated by the optional DOM marking with animal objects), providing further support for a preferred strategy favoring subject relative clauses. DOM modulates this preference, albeit to a different extent, depending on the group.
When considering the entire range of scrutinized RCs, after comparing the time-course of processing between the two conditions (each with two levels) (see Figures A1 and A2 in Appendix A), we found that object RCs with human postverbal subjects emerged as the most demanding at the critical region (“La gorila que abrazó la biológ...”) in both groups. Our participants seemed to apply an SVO default strategy, as suggested from the results of the RC interpretation questions (about 30% of our monolinguals surprisingly interpreted these postverbal subjects as objects and our bilinguals interpreted them nearly by chance, with results of 54.55%; see Table 2). However, bilinguals did not incur processing penalties as they easily accommodated the object RC reading due to the (extended) absence of DOM in their other language, Catalan. The absence of the ‘a’ marker does not seems to disrupt or distort interpretation in bilingual speakers to the same extent as in monolingual speakers. Bilinguals, particularly those living in bilingual social environments (Meisel 2021), are better-equipped to integrate a variety of objects due to encountering non-marked objects with greater frequency and in a wider range of the animacy scale in the other language at issue. The misinterpretation also suggests that bilinguals tend to adopt the well-known unmarked strategy by initially identifying an RC as a subject RC. More evidence supporting this strategy arises from the clear preference of our two groups to interpret (ambiguous) RCs with postverbal animal elements (9a) as subject SVO RCs, with only 18.92% of monolinguals and 10.81% of bilinguals, respectively, interpreting them as object OVS RCs (see Table 2).

Regarding our prediction under RQ2, analyses of the two data sources, namely, the RC interpretive questions and processing results, show that both our monolingual and bilingual participants followed a subject RC strategy. Across sentences containing postverbal human NPs, object RCs exhibited lower accuracy scores and were processed more slowly than subject RCs, consistent with findings from previous studies on Spanish (Betancort et al. 2009; Del Río et al. 2012; Perpiñán and Moreno Villamar 2013; Sagarra et al. 2020) as well as those reported for other languages (King and Just 1991 for English; Holmes and O’Regan 1981 for French; Frazier 1987 for Dutch; etc.). From the results of the interpretive questions, we determined that object RCs were (mis)interpreted as subject RCs, with participants applying the easiest SVO strategy whenever possible. The application of this strategy resulted in two outcomes, as follows: (1) a plausible interpretation in the case of permanently ambiguous OVS sentences, i.e., where the postverbal element is an unmarked animal and is interpreted as a postverbal subject, and (2) an anomalous interpretation in the case of postverbal unmarked human elements, which were more commonly accepted as objects by the bilingual group due to the smaller DOM extension in Catalan. These findings suggest an interaction between the type of RC and DOM. As mentioned earlier, the preference of subject RCs over object RCs is modulated by animacy (Betancort et al. 2009). However, this study contrasted both animate and inanimate subjects and objects within RCs, making it unclear whether this preference was due to animacy or DOM, given that animacy and DOM overlap (animate NPs have DOM, but inanimate NPs do not). In our case, we did not mix animacy and DOM; therefore, we suggest that Betancort et al.’s findings are also due to DOM rather than animacy alone. Incorporating only animate nouns that differ in their [+/- human] character can provide a more fine-grained picture of how semantic information and, specifically, obligatory/optional DOM marking can affect RC processing.

Our finding regarding the taxing processing of object OVS RCs with human postverbal subject NPs seems to stem from the application of a default subject RC strategy, prompting participants to initially interpret OVS sentences as SVO. We suggest that this leads to a subsequent reanalysis once our participants realize the absence of DOM (i.e., the postverbal subject). This reanalysis is particularly pronounced in the case of human postverbal elements, where DOM is expected to align with the initial SVO analysis. On the other hand, postverbal non-human elements are less cognitively taxing because they can easily fit into an SVO analysis, given that they are optionally marked and can be interpreted as true objects, a conclusion supported by our interpretative data (or as subjects, albeit to a much lesser extent). It is precisely in OVS sentences that different processing patterns are
observed between monolinguals and bilinguals. When processing OVS relative clauses (with postverbal human nouns), our bilinguals did not seem to be particularly taxed. This may suggest that bilinguals do not encounter too many difficulties integrating postverbal nouns without DOM as objects, as they typically do this in their other language, Catalan. An interaction appears in the critical region among bilinguals (Table 7) but not among monolinguals (Table 5). However, it should be noted that this observation is not statistically supported, as no inferential statistics have been conducted between groups. This was not the primary aim of our study. Overall, this finding underscores the general pattern favoring subject over object RCs.

There is still an issue that deserves some attention. Contrary to what is generally discussed in bilingual processing, bilinguals do not seem to be cognitively taxed when processing RCs under different conditions, except for object RCs with human postverbal nouns in the final wrap-up region (compare Figures A1 and A2). The hypothesis that bilingual processing is less automatic and less affected by cognitive functions (Bialystok et al. 2008; Costa et al. 2006), particularly when complex structures are involved (Marian and Fausey 2006), does not seem to fit very well in our case. The bilingual processing of RCs does not appear to be less automatic compared to monolingual processing, a finding consistent with earlier studies involving highly balanced bilingual speakers (Sagarra et al. 2017). As shown earlier, the current study includes bilingual university students with full competence in both languages and with excellent language and communication skills, as assessed by self-reported assessments. However, it is important to recognize that relying on self-reported measures poses a limitation. In addition, they are continuously exposed to their two languages in diverse sociolinguistic contexts and engaged in bilingual conversations. In addition, we controlled for the early age of onset of bilingualism. Thus, we have a cohort of highly balanced and proficient bilinguals. These sociolinguistic factors, together with language proficiency, have been shown to significantly influence executive functioning in bilingual individuals (Costa et al. 2006). It remains plausible that cognitive control plays a role in the interpretation and processing of complex structures such as RCs involving DOM. However, we suggest that our bilingual participants did not seem to be as cognitively taxed, as their cognitive control may be higher and more efficient compared to unbalanced bilinguals. For this reason, it would be very interesting to test different degrees of dominance to elucidate the explanatory role of language dominance and language exposure in cognitive control when processing complex structures like relative clauses.

Author Contributions: Conceptualization and methodology, A.B.; Investigation, A.B.; Data curation, A.B. and R.B.; Writing—original draft, A.B. and R.B.; Writing—review & editing, A.B. and R.B.; Funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Agencia Estatal de Investigación (AEI), Spain, grant number #FFI2016-75082-P and #PID2020-114276GB-I00.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Committee for Ethical Review of Projects at Pompeu Fabra University (CIRES-UPF, reference code 0032).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.
Appendix A

Figure A1. Reading times (in ms) in each word region for Spanish monolinguals.

Figure A2. Reading times (in ms) in each word region for Catalan–Spanish bilinguals.

Appendix B

Questionnaire of linguistic uses (for Catalan–Spanish bilingual speakers)

Edad (Age): Lugar de nacimiento (Place of birth):
Lugar de residencia actual (Current place of residence):
1. Si no es donde naciste, indica desde cuándo vives en el lugar actual (If it is not where you were born, indicate how long you have lived in the current place):
2. Lugar de nacimiento del padre (Father’s place of birth):
3. Lugar de nacimiento de la madre (Mother’s place of birth):
4. Indica a qué edad comenzaste a escuchar de forma continuada el castellano (at what age you began to listen to Spanish continuously):
5. Indica a qué edad comenzaste a utilizar (hablar) el castellano: (at what age you started to use (speak) Spanish)
6. Indica cómo y/o dónde aprendiste el castellano: (how and/or where you learned Spanish)
7. Indica a qué edad comenzaste a escuchar de forma continuada el catalán: (at what age did you begin to listen to Catalan continuously):
8. Indica a qué edad comenzaste a utilizar (hablar) el catalán: (at what age you started to use (speak) Catalan)
9. Indica la lengua (castellano, ambas, otras) que habitualmente utilizas para hablar con: (Indicate the language (Spanish, both, others) that you usually use to speak with)
   padre (father): madre (mother): hermano/as (siblings): novio/a (boyfriend/girlfriend):
10. Si de pequeño hablabas con tus padres o hermanos en alguna otra lengua de la que utilizas actualmente, indica a qué edad se produjo el cambio (If when you were little you spoke with your parents or siblings in a language other than the one you currently use, indicate at what age the change occurred):
   padre (father): madre (mother): hermano/as (siblings): novio/a (boyfriend/girlfriend):
11. ¿Qué otras lenguas puedes utilizar/utilizas normalmente (hablar, leer, escribir)? (What other languages can you/do you normally use (speak, read, write)?)
12. ¿A qué edad iniciaste el aprendizaje formal de estas lenguas? (At what age did you start formally learning these languages?)
13. Señala con un círculo la opción que mejor te representa en cada una de las siguientes preguntas (Please circle the option that best represents you for each of the following questions):
   a. ¿Qué nivel de comprensión oral tienes en cada una de estas lenguas? (What level of oral comprehension do you have in each of these languages?)
      English: perfectly good sufficient very little
      French: perfectly good sufficient very little
      Spanish: perfectly good sufficient very little
      Catalan: perfectly good sufficient very little
   b. ¿Qué nivel de comprensión lectora tienes en cada una de estas lenguas? (What level of reading comprehension do you have in each of these languages?)
   c. ¿Qué fluidez tienes en cada una de estas lenguas? (How fluent are you in each of these languages?)
   d. ¿Cómo es tu pronunciación en cada una de estas lenguas? (How is your pronunciation in each of these languages?)
   e. ¿Cómo escribes en cada una de estas lenguas? (How do you write in each of these languages?)
In response to the following questions (1–4), indicate the frequency of use of Catalan and Spanish at different ages and in the different situations proposed. To do this, use the following scale and circle the corresponding option:
   1 = Sólo catalán—Only Catalan
   2 = Fundamentalmente catalán—Mainly Catalan
   3 = Más catalán que castellano—More Catalan than Spanish
   4 = Más castellano que catalán—More Spanish than Catalan
   5 = Fundamentalmente castellano—Mainly Spanish
   6 = Sólo castellano—Only Spanish
14. ¿Siendo un niño pequeño, antes de iniciar la etapa escolar? (Being a small child, before starting the school stage?)
   Sólo catalán—Only Catalan     Sólo castellano—Only Spanish
   1 2 3 4 5 6
15. ¿Siendo un niño, en la etapa de la educación primaria? (Being a child, in the stage of primary education?)
   EN LA ESCUELA—At school
   Sólo catalán—Only Catalan     Sólo castellano—Only Spanish
   1 2 3 4 5 6
   EN CASA—At home
   EN OTROS LUGARES—Other places
16. ¿En la pubertad, en la etapa de la educación secundaria y el bachillerato? (At puberty, at the stage of secondary education and high school?)
   EN LA ESCUELA—At school
   EN CASA—At home
   EN OTROS LUGARES—Other places
17. ¿En la edad adulta? (In adulthood?)
   EN LA UNIVERSIDAD O EN EL TRABAJO—At University or at work
   EN CASA—At home
   EN OTROS LUGARES—Other places
The assessment of access to the non-standard Catalan variety and exposure was not directly conducted via a questionnaire, as pointed out by an anonymous reviewer. However, our confidence in the knowledge and usage of this variety by our bilingual speakers stems from our extensive understanding of the Catalan context and its linguistic nuances, as well as from recent findings by Benito (2023). Moreover, a recent study by Pineda (2023) further supports the prevalence of this phenomenon across all Catalan dialects, involving approximately 400 speakers. This study used individual interviews and two tasks, namely sentence production and grammaticality judgments, demonstrating that the use of ‘a’ with animate (non-)definite objects is widespread across dialects.

We also conducted parallel statistical analyses that included Group as a between-subject fixed factor. However, no significant results involving Group (either alone or in interaction) emerged from this comprehensive analysis. Moreover, the effects of other factors within this overarching analysis appeared blurred or seemed annihilated, resulting in the emergence of perplexing patterns. In contrast, conducting separate analyses to examine each subpopulation individually revealed clearer patterns within each group.

A note regarding the use of raw RTs instead of residual RTs is needed. We calculated residual readings times (adopting Trueswell and Tanenhau’s 1994 transformation procedure). We conducted analyses with both raw and residual RTs and the patterns of significance were parallel in both analyses. In this situation it is usual to report only one set of analyses (Keating and Jegerski 2015, p. 22). Therefore, we present results with raw RT data, as they are clearer.

Albeit our sentence items regarding this condition were linguistically ambiguous, as explained in relation to example (9b) above, they were not truly ambiguous for our participants since they widely selected the interpretation compatible with the subject RC. Our interpretation for this lack of ambiguity in interpretation is just that the easier subject RC strategy and the absence of DOM (DOM is optional in Spanish with animal objects) interfere in interpretation giving rise to the preferential reading attested.

As mentioned earlier in this text, we cannot specifically address Putnam et al.’s (2018) proposal that CLI predominantly flows from the dominant language to the non-dominant one since our study focuses solely on a group of balanced bilingual speakers. Recently Benito (2023) has conducted a bidirectional study with more than a hundred of Catalan–Spanish bilinguals, from a similar origin to the participants in our current research, but with different degrees of language dominance covering a spectrum of bilingualism with varying degrees of relative language dominance, as measured by the Bilingual Language Profile (BLP, Birdsong et al. 2012). The analysis of DOM acceptability and production in Catalan reveals that CLI is indeed modulated by language dominance. Specifically, the stronger a bilingual’s dominance in Spanish, the greater the influence of Spanish on Catalan DOM. While there is observable influence from Catalan in bilingual Spanish, it is not as evident in the opposite direction (specially in acceptability), indicating an asymmetry in directionality. The lack of opposite bidirectional effects partially aligns with the predictions derived from Putnam et al. (2018), suggesting that factors beyond language dominance also play a role in explaining the source of CLI.

### References


Betancort, Moisés. 2006. El procesamiento de oraciones de relativo de sujeto vs. objeto en español. *Cognitiva* 18: 55–73. [CrossRef]


Mak, Pim, Wietse Vonk, and Herbert Schriefers. 2006. Animacy in processing relative clauses: The hikers that rocks crush. Journal of Memory and Language 54: 466–90. [CrossRef]


Thane, Patrick D. 2024a. Acquiring differential object marking in heritage Spanish: Late childhood to adulthood. *International Journal of Bilingualism*. First published online February 25. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and / or the editor(s). MDPI and / or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.